Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: https://orca.cardiff.ac.uk/id/eprint/103793/

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Kuster, Corentin, Rezgui, Yacine ORCID: https://orcid.org/0000-0002-5711-8400 and Mourshed, Monjur ORCID: https://orcid.org/0000-0001-8347-1366 2017. Electrical load forecasting models: a critical systematic review. Sustainable Cities and Society 35 , pp. 257-270. 10.1016/j.scs.2017.08.009 file

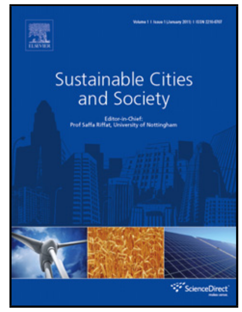
Publishers page: http://dx.doi.org/10.1016/j.scs.2017.08.009 <http://dx.doi.org/10.1016/j.scs.2017.08.009>

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See
http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

information services
gwasanaethau gwybodaeth

Accepted Manuscript



Title: Electrical load forecasting models: a critical systematic review

Authors: Corentin Kuster, Yacine Rezgui, Monjur Mourshed

PII: S2210-6707(17)30589-9
DOI: http://dx.doi.org/doi:10.1016/j.scs.2017.08.009
Reference: SCS 728

To appear in:

Received date: 9-6-2017
Revised date: 13-7-2017
Accepted date: 7-8-2017

Please cite this article as: Kuster, Corentin., Rezgui, Yacine., & Mourshed, Monjur., Electrical load forecasting models: a critical systematic review.*Sustainable Cities and Society* http://dx.doi.org/10.1016/j.scs.2017.08.009

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Electrical load forecasting models: a critical systematic review.

Corentin Kuster, Yacine Rezgui, Monjur Mourshed

BRE Trust Centre for Sustainable Engineering, School of Engineering, The Parade, Cardiff University, Cardiff CF24 3AA, United Kingdom

KusterC@cardiff.ac.uk, RezguiY@cardiff.ac.uk, MourshedM@cardiff.ac.uk

Highlights

- A systematic review protocol provides unbiased and meaningful meta-information
- A direct model accuracy comparison across studies is meaningless
- A taxonomy for an informed forecasting model's selection is proposed
- Recommendations on writing electrical load forecasting related paper are given

ABSTRACT

Electricity forecasting is an essential component of smart grid, which has attracted increasing academic interest. Forecasting enables informed and efficient responses for electricity demand. However, various forecasting models exist making it difficult for inexperienced researchers to make an informed model selection. This paper presents a systematic review of forecasting models with the main purpose of identifying which model is best suited for a particular case or scenario. Over 113 different case studies reported across 41 academic papers have been used for the comparison. The timeframe, inputs, outputs, scale, data sample

ACCEPTED MANUSCRIPT

size, error type and value have been taken into account as criteria for the comparison. The review reveals that despite the relative simplicity of all reviewed models, the regression and/or multiple regression are still widely used and efficient for long and very long-term prediction. For short and very short-term prediction, machine-learning algorithms such as artificial neural networks, support vector machines, and time series analysis (including Autoregressive Integrated Moving Average (ARIMA) and the Autoregressive Moving Average (ARMA)) are favoured. The most widely employed independent variables are the building and occupancy characteristics and environmental data, especially in the machine learning models. In many cases, time series analysis and regressions rely on electricity historical data only, without the introduction of exogenous variables. Overall, if the singularity of the different cases made the comparison difficult, some trends are clearly identifiable. Considering the large amount of use cases studied, the meta-analysis of the references led to the identification of best practices within the expert community in relation to forecasting use for electricity consumption and power load prediction. Therefore, from the findings of the meta-analysis, a taxonomy has been defined in order to help researchers make an informed decision and choose the right model for their problem (long or short term, low or high resolution, building to country level).

KEYWORDS

Electric consumption and load prediction; forecasting models; Machine Learning; Regression; Time Series Analysis; Long-term/short-term forecasting

1 INTRODUCTION

Forecasting models are widely used in different domains; e.g. in finance to forecast stock exchange courses or indices of stock markets (Bianco et al. 2009), in business to schedule

ACCEPTED MANUSCRIPT

staff, manage inventory and predict demand (Hyndman & Athanasopoulos 2014), in medicine to monitor the spread of diseases (Generous et al. 2014), and in meteorology for predicting weather. Equally, forecasts play an essential role in the control of power plants and electric power exchange in interconnected systems (Mohandes 2002). Forecasting supports energy planners in understanding the influence of some variables on energy consumption and thus inform decision making (Al-Ghandoor et al. 2009). On a temporal scale, forecasts can be *short-term* for instance for balancing electricity supply; and *long-term* forecasts, including for capacity expansion, capital investment return studies, and revenue analysis (Parlos et al. 1996). Over the years, many different forecasting models have been applied for electricity and power predictions such as multivariate and multiple regression (Azadeh & Faiz 2011; Filik et al. 2011; Wang 2012; Al-Hamadi & Soliman 2005; Farzana et al. 2014), SVM (Massana et al. 2015; Mohandes 2002; Garulli et al. 2015), time series (including Autoregressive Integrated Moving Average (ARIMA) and the Autoregressive Moving Average (ARMA)) (Fan et al. 1994; Gonzales Chavez et al. 1999; Hoffman 1998; Chujai et al. 2013). Equally, artificial neural networks (ANN) have become widely used for prediction scenarios. ANN has been used for various tasks such as (a) short-term load forecasting (STLF) in microgrids (Hernandez et al. 2014; Hernández et al. 2014; Hernandez et al. 2013; Hernández et al. 2012; Twanabasu & Bremdal 2013), (b) optimisation scenarios at building level (Platon et al. 2015; Hsiao 2015), and (c) long term horizon scenarios to determine annual electricity consumption of a region, district or building (Farzana et al. 2014; Azadeh & Faiz 2011). There is no consensus over a particular forecasting model and the use of a method over another is often the result of the expert's preference. Moreover, in his paper showing the outcomes of the M3 competition (competition comparing the accuracy of different forecasting methods, realised in 1982, 1993 and 2000), Makridakis states that "simple methods developed by practicing forecasters do as well, or in many cases better, than

ACCEPTED MANUSCRIPT

sophisticated ones" (Makridakis & Hibon 2000). This means that there are no evidences that complex models will outperform "simple" ones. Therefore, it is relevant to identify which model fits a particular situation.

The need for forecasting varies from one scenario to another; the setting of a model is subject to numerous variations: including the available data used as inputs, the timeframe wanted, the time resolution (from every minute to annually), the scale (from a simple building to a whole country consumption). The aim of this review is to critically analyse and identify the quality of a method compared to some other potential solutions in a specific forecasting scenario and to assist users in their forecasting method's selection by simply answering questions such as "*Which model do I need to generate hourly electricity demand/consumption of a building for the next 2 years?"*.

The paper will first introduce the systematic review process employed (Section 2.1) for the selection of case studies. Equally, a short description of the most commonly encountered forecasting models is given. Each of these forecasting models has advantages and disadvantages and none is 100% efficient. It is important to know their limitations before considering their use. Section 3 gives an overview of the main context characteristics of the study cases across the paper references (e.g. location, year, scale, data used, model used, timeframe considered…) as well as the results of their application in various scenarios and field studies. A taxonomy for the decision making of prediction models is presented as the main output of this review. Finally, a discussion section elaborates on the open questions resulting from the study and suggestions of the authors.

2 METHODOLOGY

The study is based on the critical review of academic research aimed at power and electricity forecasting. The selection of the different papers has followed a rigorous systematic protocol.

ACCEPTED MANUSCRIPT

In this section the systematic process used for the review is described. The different steps for the papers' selection are described. The different keywords used and domains' restrictions are explained for an objective, non-biased papers' selection. Additionally, some of the most popular forecasting models namely ANN, Time series analysis (including AR, MA, ARMA, ARIMA, SARIMA), SVM and Bottom-up model are briefly explained in this section.

2.1 *Systematic review protocol*

For the study, a systemic approach of the literature has been employed. Systematic reviews vary from the traditional review by extensive literature searches and meta-analysis of the finding, reducing the effect of chance and biases (Tranfield et al. 2003). A systematic review must follow a well-defined protocol introduced to bring more clarity, rigour and repeatability. The author must first define the research question(s); then define the research criteria to apply in order to select accurate publications. Once the selection done, the author can analyse the data and finally discuss the results (Righi et al. 2015; Higgins JPT, Green S 2006). In this process, the selection of the criteria is particularly important. The research criteria have been selected according to the research question. While developing the research question some keywords appeared naturally like "electricity forecasting models", "electricity prediction models" or "electricity demand models". Using the online database Scopus, one of the established abstract and citation databases of reviewed literature (Anon n.d.), and its advanced search tool, the search results were first limited to these keywords. The appearance of the keywords in the main text has been excluded as criteria due to a high probability of occurrences. The threefold Title-Abstract-Keyword provides more relevant results because it targets better the global content of the text. On this 1st search, 10 667 results were returned. A statistical analysis shows the distribution of the returned results by area of study. From the 10 667 results 39.1% are from Engineering; 38.1% are from Energy; 16.3% from Computer

ACCEPTED MANUSCRIPT

Science and 14.6% from Environment science. The other areas are not relevant in this study and therefore have been excluded (Figure *1*).

Texts in languages other than English have been excluded. The reduction to the four areas: Engineering, Energy, Computer science and Environmental Science leads to a new selection of 5845 papers. In order to fit even better to the desired topic, other keywords were targeted within the previous results. The texts including the keywords “building”, “dwelling” or “household” inside the title, abstract or keywords have been selected. 900 papers have been identified under these criteria. Finally, Scopus provides the overall of the keywords allocated to each paper. A quick overview on the keywords has enabled the identification of some irrelevant papers like “electric vehicles”, “wind power” or “global warming”. In order to avoid irrelevant studies, papers associated to the specific keywords “electricity demands” and “electric load forecasting” have been selected. At the end, 153 have been identified and will constitute the study basis. The whole process is shown in Figure 2

Among the 153 studies, 76 are articles, 68 are conference papers, five are reviews, three are articles found in the press and one is a short survey; all from 44 different countries. Among the 153 references found on the topic, 41[1] have been reviewed in depth. Within the 41 references, 113 different implementations of forecasting models have been identified. Having explicit criteria against which to assess studies helps to avoid hidden bias, by having clear consistent rules about which studies are being used to answer the review's specific research

[1] (Abdel-aal & Al-Garni 1997; Al-Ghandoor et al. 2009; Al-Hamadi & Soliman 2005; Aydinalp et al. 2004; Azadeh & Faiz 2011; Beccali et al. 2008; Bianco et al. 2009; Boulaire et al. 2014; Cheng & Steemers 2011; Chujai et al. 2013; Ciabattoni et al. 2013; Dilaver & Hunt 2011; Fan et al. 2015; Farzana et al. 2014; Filik et al. 2011; Fischer et al. 2015; Garulli et al. 2015; Gonzales Chavez et al. 1999; Gul et al. 2011; Hernández et al. 2014; Hoffman 1998; Hsiao 2015; Inglesi 2010; Jurado et al. 2015; Koprinska et al. 2011; Marvuglia & Messineo 2012; Massana et al. 2015; Mathieu et al. 2011; McLoughlin et al. 2013; McLoughlin et al. 2012; Mena et al. 2014; Mohandes 2002; Newsham & Birt 2010; Platon et al. 2015; Richardson et al. 2010; Swan et al. 2011; Twanabasu & Bremdal 2013; Wang 2012; Widen & Wackelgard 2010; Yoo & Hur 2013; Zahedi et al. 2013)

ACCEPTED MANUSCRIPT

questions (Eppi 2007). Thus, the systematic review approach allows the reduction of biases and to consider the distribution of the cases as a good representation of the overall framework.

2.2 *Cases comparison*

From this point, the cases have been rigorously studied following specific characteristics. A relevant selection of a case characteristic is important because it is the starting point for an accurate and meaningful comparison between different electricity forecasting models. The idea is to best represent a particular situation through those characteristics without leaving out aspects that could influence the forecasting performance. A first analysis of the references helped in this matter. The authors have identified the different characteristics to take into account by considering the aspects that repeatedly appeared in the literature in order to describe a case. Table 1 gives the characteristics considered and a description of their suitability.

From there, a second analysis has been done and an excel spreadsheet has been populated with all the needed information for every case.

2.3 *Forecasting models highlights*

From the 113 cases studied, 16 different models have been identified. The first observation is that some models can be categorised under a same label. For example, AR, MA, ARMA, ARIMA, seasonal or not, with or without exogenous variables can be seen as a part of time series analysis model. Therefore, they will be gathered into the label “Time Series”.

It is interesting to study the models’ distribution through all the references in order to have a representation of the current trend in the forecasting model use. If the models’ distribution does not give an exact representation of the practices of the expert community, it still

ACCEPTED MANUSCRIPT

provides a good overview. Figure 3 illustrates the distribution of the different analysed forecasting models. Because one paper can proceed to several applications on one specific prediction model, the distribution of each forecasting model through the reviewed papers provides a better representation of the actual trend. Therefore, distribution will always be considered as a number of papers in which a particular forecasting model is used rather than the number of application.

A clear trend is observed in the use of forecasting models. The regression model (most often multiple regressions or multivariate regressions) is the most widely used, and is present in 17 papers out of 41 (43.6% of the papers), followed by the artificial neural networks (ANN) present in 15 papers (38.5%). Time series models are present in 30.8% of the papers, i.e. 12 papers. In a lesser proportion, SVM and Bottom up models are used in 15.4% and 10.3%, respectively. The other models are singularities. The relatively high quantity of regression, ANN and time series models can be explained by their popularity in the research community. This observation strengthens their status as leading models in the field. SVM and bottom up model are present in a lesser extent but there is a clear framework developed around these models sustained by an increasing number of studies. Overall, five main models were identified from the review of the articles. A short description of the most encountered forecasting models is given below.

2.3.1 *Artificial Neural Network*

Conventional models such as regression are limited and can sometime lead to unsatisfactory solutions (Aggarwal & Song 1997). The reasons include the too high number of computational possibilities leading to large solution times and the complexity of certain non-linear data patterns (Aggarwal & Song 1997). On this type of challenges, artificial neural

ACCEPTED MANUSCRIPT

networks and intelligent machine learning technique, provide a promising and attractive alternative. The increasing computational power has facilitated forecasting in a large set of power system management from load forecasting to security assessment or fault diagnosis (Wehenkel 1997). However, on some problems, the use of conventional models lead to unsatisfactory solutions due to the high complexity of variables' relationships and the extent of computation power requirements (Landau & Taylor 1998). It is in these cases that artificial neural networks (ANN) are used. Two references have been mainly used for the ANN description: the work of Raj Aggarwal and Yonghua Song that gives an introduction to the field of ANN via three tutorials which are proposed to engineers with an application in power systems (Aggarwal & Song 1997; Song & Aggarwal 1998a; Song & Aggarwal 1998b); and the book of Lawrence Jay Landau and John Gerald Taylor that gives a broad view on the concept of neural networks. It explains the basics of artificial neural networks and the mathematical underpinnings (Landau & Taylor 1998). ANN is an intelligent machine learning method based on the structure of the human brain. As the human brain, ANN is composed of neurons and interactions within multiple layers. Even if current ANNs are far from reflecting the complexity of a human brain, they remain powerful tools in pattern recognition. A neuron is the main element of the ANN, it can receive or send a normalised signal from and to the other neurons of the network. The wires between neurons are called “weight” wkp, one for each wire coming to a neuron from another one. Overall, there are three main features that determine an ANN: the architecture of the net (feedforward or recurrent), the learning rule used for defining the weights during training (perceptron, Hebbian, etc.), the activation function between neuron input and output. One of the most commonly used ANN is the Multi-Layer-Perceptron. This multilayer network is based on a backpropagation rule which evaluates the output’s error and reduces it, adjusting the weights by back-propagating the error from the output to the hidden layer. ANNs are particularly

ACCEPTED MANUSCRIPT

suited for energy forecast. They provide a good estimation in cases where data is incomplete (Aggarwal & Song 1997), and can address complex nonlinear problems while demonstrating robustness and fault tolerance (Zhai 2005). More, it is a data-driven self-adaptive model (Zhai 2005; Pantic 2000) that (a) includes pattern recognition and captures subtle relationships (Aggarwal & Song 1997; Pantic 2000), (b) deals with noise (Aggarwal & Song 1997), (c) does not depend on the programmer's prior knowledge of rules (Song, 1997); and (d) identical and independent operations can be done simultaneously (Aggarwal & Song 1997). However, ANN’s results cannot be easily explained as (a) they are not mathematically based (Aggarwal & Song 1997), (b) it is computation time consuming (Aggarwal & Song 1997), (c) the training process optimisation is complex (Askarzadeh & Rezazadeh 2013), (d) extended data is required (Zhai 2005) and (e) the model may never converge in some cases (Zhai 2005).

2.3.2 *Time series analysis*

Some of the most widely used methods for time series analysis and forecasting are the Autoregressive Integrated Moving Average (ARIMA) and the Autoregressive Moving Average (ARMA). The ARMA and ARIMA have been introduced in 1970 by two statisticians, George Box and Gwilym Jenkins (Box et al. 2008). The basic ARMA model is composed of an autoregressive model (AR) and a moving average model (MA). The autoregressive model is a linear regression of the current value based on one or more previous values. Just as an AR, the MA is a linear regression, at the difference that it regresses current values against the white noise or errors of one or more past values. Note that an essential condition to process an ARMA model is that the time series is stationary. If not, the stationarity is achieved by differencing a non-stationary series in first place. The introduction of this step lead to a new model called ARIMA with the “I” standing for “Integrated”. In order to deal with the seasonality, Box and Jenkins introduced a new model,

ACCEPTED MANUSCRIPT

the seasonal ARIMA or SARIMA. The most commonly used seasonal ARIMA is probably the ARIMA(0,1,1)x(0,1,1) which corresponds to a seasonal exponential smoothing model. Overall, Box-Jenkins forecasting model is (a) adaptable, (b) can deal with seasonality and with non-stationarity and (c) only requires the past value of a time series (Zhai 2005). Nevertheless, it is unlikely to perform well on long-term prediction (Zhai, 2005), is computation time-consuming (Zhai 2005), is subjective and requires a good understanding of the underlying statistics (Zhai 2005) .

2.3.3 *Bottom up end-use approach*

We call "bottom up approach" the construction of a complex system by aggregating elementary systems. Applied to the electricity consumption, it is simply the aggregation of all appliances loads within a household in order to determine the overall load of this household. The bottom up approach the most commonly cited is the Capasso bottom up model (Capasso et al. 1994). This approach evaluates the probability for a specific appliance to be "on" at every time step of a day by considering various factors involving the appliances and household members' characteristics. Each appliance is related to one or more activities. The probability that an appliance is "on" is then linked with the probability for an activity to be done at a certain time of the day by one or more members. A calibration is applied to this probability taking into account (i) if the activity can be done by the person available, (ii) how many people the activity requires, and (iii) can the activity be done simultaneously with another activity. Once the activity probability is computed, socio-economic criteria are used to determine the penetration of appliances in the household. The power requirement and duration of each of these appliances is extracted and used to determine if they are suitable for a given activity and building type. For that, the minimum duration of usage of an appliance has to fit the activity probability in which the appliance is involved and the power required has to be smaller than P_{limit}, the maximum power load allocated to the household

ACCEPTED MANUSCRIPT

(considering that other appliances might be used in the exact same time). These steps lead to the creation of the appliance load profile and finally the daily load profile of the household. The bottom up model has the advantage to consider behaviour of the various types of customer and lifestyle-related psychological factors. It describes interrelations between appliances and members of the household and is easily understandable. Moreover, it can deal with missing values and its maintenance is simple. Among its disadvantages are the large number of data and tenants behaviour surveys required, the lack of information regarding customers' behaviours in the long-term, thus inherently inaccurate in the long-term (Ghods & Kalantar 2011) and that the model assumes a constant relationship between electricity consumption and end-use (Ghods & Kalantar 2011) .

2.3.4 *Support Vector Machine*

Support Vector Machines have been first introduced by Vladimir Vapnik with a paper at the COLT 1992 conference (Boser et al. 1992). Then, in 1995, the soft margin classifier was introduced by Cortes and Vapnik in the paper Support Vector Networks (Cortes & Vapnik 1995). Originally, SVMs were created to deal with pattern classification problems like character recognition, face identification and text classification. In 1995, Vladimir Vapnik extends SVM to a regression algorithm in his book, The Nature of Statistical Learning Theory (Cherkassky 1997). Over the years various applications were found in the literature; e.g. time series prediction problem. The purpose of an SVM is to create an optimal separating hyperplane in a higher dimensional feature space such that subsequent observations can be classified into separate subsets. In practice, real data are not as perfectly separable. In order to provide a hyperplane, one has to relax the requirement that a separating hyperplane will perfectly separate every training observation. For that, a soft margin classifier (SVC) has been constructed. In the case of non-linear boundaries, the use of SVM is convenient (Auria & Moro 2008). Indeed, the SVM allows non-linear decision boundaries by using an

ACCEPTED MANUSCRIPT

appropriate transformation that makes them linear on a higher dimensional feature space. Unfortunately, computation on high dimension feature space can be very costly and SVM depend a lot on the proper selection of the hyper-parameters (Adhikari & Agrawal 2013). To improve the computation efficiency, a solution also called the “Kernel trick” is used (Adhikari & Agrawal 2013). Kernels are functions used to represent inner products between observations rather than observations themselves. Thus, it modifies how we calculate "similarity" between two observations in a more flexible way, allowing to change and solve a non-linear problem by a linear problem on a higher-dimensional space.

2.3.5 *Regression*

A regression is the simple statistical method that allows the observation of relationship between variables. Thus, the response, outcome, or dependent variable can be defined by other variables called predictor, explanatory, or independent variables. The most common form of regressions analysis used for prediction are the linear regressions and the polynomial regressions. The linear regression links the response y and the predictor x by the simple linear model:

$$y = \beta_0 + \beta_1 x + \varepsilon.$$

Where $\beta 0, \beta 1$ and ε are the intercept and the slope of the line and the random “error” respectively (Hyndman & Athanasopoulos 2014). The extension of the simple linear regression is the multiple linear regression. The difference between simple and multiple linear regression being the number of variables introduced as predictors that goes from one variable in the simple model to several in the multiple. Thus, the regression can not only be time related but also integrate some other independent variables. Four conditions, however, must be taken into account: the mean of the response at each set of values of the predictors is a linear function of the predictors, errors are independent, errors at each set of values of the

ACCEPTED MANUSCRIPT

predictors are normally distributed, errors at each set of values of the predictors have equal variances (Anon n.d.). In the same way, the polynomial regression is a regression analysis where the predictor is related to the response via a polynomial of degree n. It is used to fit nonlinear data.

3 FORECASTING MODELS COMPARATIVE ANALYSIS

The initial phase involves analysing each paper's scope and scenario objectives. The focus is then on the analysis of the forecasting models used, prediction horizon, variables and processes employed. Finally, key patterns in the use of the selected forecasting models are described.

In term of application, the objectives are various. The most frequently encountered objective is the demand response for production and distribution of electricity. This objective requires short-term horizon predictions with high-resolution data in order to have a fast response to the electrical loads. This can be applied to a single building when several electricity sources are involved (Mathieu et al. 2011; Mena et al. 2014) or at district level with the increasing development of smart grids (Garulli et al. 2015; Hernandez et al. 2014; Hernández et al. 2014). Another application is associated with mid to long term forecasting (1 week to a year) where the prediction is employed for power system planning (Al-Hamadi & Soliman 2005), maintenance or production and resell market (Filik et al. 2011). Long term prediction (several years) with large time step are often applied for policy making (Azadeh & Faiz 2011), large scale planning (Gonzales Chavez et al. 1999), statistical prevision (Bianco et al. 2009) or business plan (Wang 2012). Some applications are more household oriented with a great emphasis on appliances' running time and on occupancy (Ciabattoni et al. 2013; Fischer et al. 2015; Widen & Wackelgard 2010; Richardson et al. 2010). In the latter cases, the researcher seeks a reproducible model that can fit several household types. Thus, the prediction can

ACCEPTED MANUSCRIPT

apply from a single house up to an entire district by aggregation. Finally, some studies attempt to cover all the previously cited applications with high a resolution model that has a long-term horizon (Filik et al. 2011).

3.1 *Scope and scenario objectives*

This section presents the different use cases found in the literature by giving their main features such as if they have been pre-processed, the timeframe, inputs, and resolution. This gives a better insight of the references listed.

3.1.1 *Forecasting data pre-processing*

Many studies have concerns about the variables they present as inputs of their models. Indeed, many studies support the use of data pre-processing in order to improve forecasting accuracy ((Chujai et al. 2013; Hsiao 2015; Azadeh & Faiz 2011)), especially when using machine learning algorithms (Crone et al. 2006; Huang et al. 2015; Suhartono & Subanar 2006). It appeared than in 66.0% of the cases, a pre-processing has clearly been done. Note that the remaining 33% do not necessarly mean a lack of pre-process but simply that it has not been mentioned. Overall, data pre-processing is a common practice for forecasting. Four different kinds of pre-processing can be identified: (1) smoothing and filling missing values, (2) measurement of variables dependency and significance, (3) data decomposition and classification and (4) check order of integration and stationarity. For that, several mathematical and statistical tools are used, the most widely spread are principal component analysis (PCA), which uses principles to transforms a number of possibly correlated variables into a smaller number of variables called principal components, Pearson correlation (PCC) which show the interdependency of sets of variables, p-value that is used for testing a

ACCEPTED MANUSCRIPT

statistical hypothesis, analysis of variance (ANOVA), Kernel density estimation (KDE) which is a non-parametric density estimator and Canonical Correspondence Analysis (CCA).

3.1.2 *Forecasting timeframe and resolution*

Another important characteristic is the time-term considered for the prediction. Indeed, the timeframe and resolution that are chosen for a prediction will highly influence the results and the choice of a model over another. The timeframe has been classified into 4 categories such as “very short term” (less than an hour), “short term” (1 hour to several days), “mid-term” (1 month to a season) and “long term” (a year or more). The time resolution represents the time-step considered for the prediction. It goes from every minutes to annually. Table illustrates the term distribution through the different papers and cases. The distribution in percentage is based on the number of paper in which the timeframe is used.

With respectively 61.5% and 43.6%, the long-term and short term prediction represents the actual needs for electrical loads forecasting in buildings. In the vast majority of cases, the use of forecasting model is for 1 hour, 1 day or 1 year ahead prediction. Very short-term and mid-term prediction are not highly represented within the cases. This can be explained by the needs of the industry for short term and long-term prediction. Indeed, short-term prediction has a direct application for quick electricity demand response while long term prediction are often used for prevision and strategies.

3.1.3 *Forecasting input variables*

For electricity and power prediction, a forecasting model can be implemented with a large range of inputs. Independent variables such as income, occupancy, electricity price,

ACCEPTED MANUSCRIPT

temperature, building size, rainfall, dwelling type, GDP, population are just few examples of the various possible inputs. For the purpose of the study, the different exogenous variables have been classified into 4 categories: "Socio-economic" related to the socio-economic situation of the zone considered, "Environmental" related to the weather conditions, "building and occupancy" related to the building type and activity and "time index". Table presents the input used distribution across the papers and cases. The building characteristics are, with 48.7% of paper considering them, the most used exogenous variables. Environmental variables and socio-economic variables follow and are present in respectively 41.0% and 38.5% of the papers. The time index data, which are simply the date stamps series introduced as an input, are used in a lesser proportion with 28.2% of papers found. Finally, models without exogenous inputs are present in 30.8% of the papers.

In order to fully understand this distribution, it is important to look at when these inputs are used. Figure 4 illustrates how the inputs are split into the different timeframe. In the same way, Figure 5 shows the repartition of the input depending on the scale of the study. Some trend can be identified. Indeed, socio-economic variables are in a high majority of cases used for long-term prediction as well as large-scale studies (from a city level to entire country). Environmental variables are implemented in models that aim to predict short-term and small to mid-scale studies (building to district level). Overall, electricity historical data (past patterns) are equally present in short or long term, small or big scale studies. If building and occupancy inputs are almost equally used for short-term and long-term prediction, they are mainly employed for small and mid-scale level (building to district). Finally, time index are mainly introduced for short-term and small-scale prediction.

ACCEPTED MANUSCRIPT

One of the reasons for input variables selection is the meaning that the developer wants to give to his / her model. Indeed, some studies target some particular variables to highlight the relationship between them and impact for instance on electricity consumption.

Finally, the introduction of a time index in certain cases has proven improving the accuracy when time series have been clustered (Hernandez et al. 2014; Yoo & Hur 2013) or when the model highly depends on the occupancy such as in (Ciabattoni et al. 2013).

3.2 *Key observations*

The study aims to identify which forecasting model best fits particular situations and variables. Indeed, following the situation and the variables available, the use of a model will be preferred in order to obtain the best accuracy possible. This section presents the different observations done on the use of forecasting models by highlighting potential correlation between a scenario parameter such as timeframe and/or inputs variables and a particular model. The outcome is the overall representation of the practices within the expert community. Those practices are assumed as being representatives of the good use of the models.

3.2.1 *Model vs timeframe*

In this section, the models have been compared by considering the timeframe they are meant to predict. Only the five most commonly encountered models that are ANN, bottom up, time series analysis, regression and SVM are being compared. The singularity of the other models listed in Figure 3 does not allow any interpretations about the case they best fit. Figure 6 shows the number of papers in which a timeframe is considered given a certain model. The

ACCEPTED MANUSCRIPT

majority of the regression models are used for long-term prediction, one year or more. Only four on 19 configurations are against short or very short-term prediction. On the other hand, ANN is mainly used for short-term prediction with 10 papers studying this configuration. In a lesser extent, the use of ANN for a long or mid-term prediction follows with a total of 6 papers. Likewise, the time series analysis and SVM models have been mainly applied for a short and very short-term prediction with respectively 10 and 6 different references considering this configuration. The bottom up model seems to be slightly preferred for long term forecast.

The regression remains widely used for long term forecasts due to its simplicity and accuracy on this timeframe, especially when the time resolution is large as states AlRashidi in Section 2 of his paper "Long term electric load forecasting based on particle swarm optimization" (AlRashidi & EL-Naggar 2010). Short term predictions require more sophisticated models such as machine learning or ARIMA because the variables interrelationship are more complex and sensitive on this time scale (Hippert et al. 2001; Ho et al. 2002).

3.2.2 *Model vs input*

This last section presents the repartition of the inputs implemented within the model

following the model.

Figure 7 shows the number of papers in which input variables are introduced in a given model. The regression models have been mainly set up with socio-economic inputs. Note that if regressions models are using socio-economic variables, a direct correlation between the model and the variable is debatable. Indeed, regression models are widely used for long-term

ACCEPTED MANUSCRIPT

forecasts (see section 3.2.1) and long-term forecasts are correlated with the use of socio-economic variables (see section 3.1.3). ANN shows a preference for environmental, building and time index inputs. Overall, ANN has been implemented with a relatively large range of variables which indicates flexibility of the models toward the data introduced as inputs. In the case of the ANN and SVM, the relatively high amount of time index data introduced is mainly due to a possible need in order to increase their accuracy. Time series analysis is most often set up without exogenous variables. In this model, exogenous variables are introduced

The

for a better performance but are not mandatory for its proper functioning. Finally, the bottom up model systematically uses building related data such as available appliances or occupancy, due to the nature of the model itself (see section 2.3.3).

3.2.3 *Model vs output resolution*

Lastly, it is interesting to look at the time step considered for the prediction. Indeed, forecasts serve different purposes that may require a specific resolution. From less than an hour to yearly, several examples have been found in the literature. Often a model is test on different output resolution since its accuracy will be relative to those. Thus, a model can show poor performance on an hourly basis but overall good accuracy on a weekly or monthly resolution.

Figure 8 shows the repartition of the different models according to the timeframe and resolution considered. If there is no evident clear trend, some observation can be done. Firstly, the graph confirms the observation done in the section 3.2.1 where ANN, SVM and time series analysis are preferred for short-term forecasts while regression and bottom up for

ACCEPTED MANUSCRIPT

long-term. Secondly, hourly resolution on short-term forecast represents slightly more cases and those in every model applied. Lastly, on long-term forecasts, regression is favoured on an annual resolution of predictions over several years while bottom up is slightly preferred on lower time steps.

Discussion

The accuracy of each studies has been investigated in order to define which model performs the best in a given scenario context. However, a direct comparison of the study cases seems irrelevant because of the numerous variables influencing their performance. Indeed, models are implemented for different locations, in different time periods, with data of more or less good quality and supported by scripts more or less well written. Even the mean for accuracy determination are different (mean absolute percentage error, mean percentage error, root mean square error (RMSE), coefficient of variance of RMSE) making comparison difficult. Overall, none of the model clearly outperforms the others and seeking the most accurate is meaningless in this case. Instead, the study assumes that the most commonly used practices by the expert community are representative of the best use of forecasting models. If the position of the author is in favour of this theory, it remains obviously debatable.

The above elements of response following the analysis of the critically reviewed papers have informed the development of a simple taxonomy summarizing the use of "major" models (ANN, regression, time series analysis, bottom up, SVM) in particular scenarios / applications. It is designed to answer the following type of questions:

"Which model do I need to generate <u>hourly</u> electricity demand/consumption of a <u>building</u> for the <u>next 2 years</u>?"

ACCEPTED MANUSCRIPT

Figure 9 presents a first taxonomy faithful to the cases found in the literature. Associated with Table , it gives real cases found in the literature in order to solve a specific problem. Thus, this taxonomy does not necessarily generalized on the model that need to be used but leads the user to references they can consult.

A broader taxonomy can be developed according to the results of the previous analysis. In the same way, the user can refer to the taxonomy and choose between the recommended models. All colours superior to a particular case can be applied to this one. Table explains the values of each colour and the model associated. For instance, the daily electricity consumption on a long term period at the building level can be forecasted using a regression or a bottom up model with building related data for instance.

It should be noted that the use of the bottom up model highly depend on the availability of precise data concerning the building(s) and their appliances.
Overall, the authors suggest that the researcher tries the few models given by the forecasting models' taxonomies fitting his or her situation (term, scale, available inputs). Thus, to the above question:

"Which model do I need to generate <u>hourly</u> electricity demand/consumption of a <u>building</u> for the <u>next 2 years</u>?"

we can answer:

"A <u>bottom up model</u> introduced with <u>high resolution and disaggregated (by appliances) historical data</u> that have been smoothed and <u>building data</u> such as occupancy, appliance

ACCEPTED MANUSCRIPT

availability etc., as exogenous inputs. Environmental data such as weather condition and time index can eventually be introduced (see Ciabattoni et al. 2013; Fischer et al. 2015).”

Last concern is about the writing of academic papers. Indeed, information retrieval had been particularly difficult and the authors suggest that while presenting his or her work, the researcher provides explicitly the frame of the forecasting implementation (timeframe, time resolution, scale, inputs, outputs, pre-processing…). Moreover, the researcher should present several means for error measurement (mean absolute percentage error, mean percentage error, root mean square error (RMSE), coefficient of variance of RMSE) in order to facilitate a direct comparison across studies.

4 CONCLUSION

In this review, 113 different applications of various forecasting models distributed into 41 international papers have been studied. Many criteria have been checked such as the scale of the project, the time-term, time resolution, input employed, data pre-processing, error etc. Overall, if the models ‘selection via a direct accuracy comparison appeared to be meaningless in this study because of external elements that can interfere, some patterns in the use of the models are interesting. Considering the numerous use cases and papers studied, it is reasonable to assume that recurrence in the use of forecasting models reflects good practices. Some models seem to be favoured for electricity and power forecasting such as multivariate regression or Multiple Linear Regression, Artificial neural network and Time series analysis. Regression models are often employed for long-term prediction where periodicity and changes are less significant. This long-term predictions are often associated with socio-economic variables and building characteristics reflecting the correlation between these variables and electricity consumption on the long–term. ANN and Time series analysis are

ACCEPTED MANUSCRIPT

mainly used for short-term predictions where electricity and power consumption patterns are more complex. Time series analysis leans principally on past electrical loads data while ANN are mainly set up with past values, Environmental and building/ occupancy data. Support vector machine and bottom up models are present in a significant amount of paper showing increasing interest thereof. In the case of SVM, they are similar to the ANN in their usage (short-term with Environmental, past values and occupancy inputs). In the case of bottom up models, they have the advantage to be easily understandable and can be used for 1 day to 1week ahead prediction at building level. However, the model requires a well detail dataset about appliances electricity consumption and occupancy. In some cases, a time series index is introduced in order to increase the accuracy of certain models. A time index is particularly useful when a model strongly depends on occupancy such as the bottom up model or when the time series has been decomposed in underlying patterns. Additionally, a pre-analysis and pre-processing of the input data is recommended in order to have better results. Indeed, it is recommended to smooth time series from errors and to fill missing values. A measurement of variables dependency and significance can help both on speed of computation and accuracy. Data decomposition and classification allow breaking down complex series into simpler models and thus give better performance on forecasting. Two simple taxonomies are presented therefore, one that leads to real cases found in the literature and a second that generalizes the outcomes of the study. When a researcher has to make a choice on the model to use, one can refer to the general taxonomy, going across the different branches of the tree that fit his or her situation and then try the recommended forecasting models. Equally, one can refer to the 1st taxonomy to have real use cases of the model coming from the literature. If these taxonomies do not have the ambition to reflect all the complexity of electricity power and consumption prediction, they nevertheless give a good overview and can lead to the selection of a potential model solution.

ACCEPTED MANUSCRIPT

Abdel-aal, R.E. & Al-Garni, A.Z., 1997. Forecasting Monthly Electric Energy Consumption in Eastern Saudi Arabia Using Univariate Time-Series. *Energy*, 22(11), pp.1059–1069.

Adhikari, R. & Agrawal, R.K., 2013. *An Introductory Study on Time Series Modeling and Forecasting*,

Aggarwal, R. & Song, Y., 1997. Artificial neural networks in power systems. Part 1: General introduction to neural computing. *Power Engineering Journal*, 11(3), pp.129–134. Available at: http://digital-library.theiet.org/content/journals/10.1049/pe_19970306.

Al-Ghandoor, A. et al., 2009. Residential past and future energy consumption: Potential savings and environmental impact. *Renewable and Sustainable Energy Reviews*, 13(6–7), pp.1262–1274.

Al-Hamadi, H.M. & Soliman, S.A., 2005. Long-term/mid-term electric load forecasting based on short-term correlation and annual growth. *Electric Power Systems Research*, 74(3), pp.353–361.

AlRashidi, M.R. & EL-Naggar, K.M., 2010. Long term electric load forecasting based on particle swarm optimization. *Applied Energy*, 87(1), pp.320–326. Available at: http://dx.doi.org/10.1016/j.apenergy.2009.04.024.

Anon, Scopus - About Scopus. Available at: https://www.elsevier.com/solutions/scopus.

Anon, STAT 501. *Department of Statistics Online Programs, Penn state college of science*. Available at: https://onlinecourses.science.psu.edu/stat501/.

Askarzadeh, A. & Rezazadeh, A., 2013. Artificial neural network training using a new efficient optimization algorithm. *Applied Soft Computing Journal*, 13(2), pp.1206–1213.

ACCEPTED MANUSCRIPT

Available at: http://dx.doi.org/10.1016/j.asoc.2012.10.023.

Auria, L. & Moro, R. a, 2008. Support Vector Machines (SVM) as a Technique for Solvency Analysis. *Discussion Papers of Deutsches Institute of Wirtschaftsforschung*, (August). Available at: http://papers.ssrn.com/sol3/JELJOUR_Results.cfm?form_name=journalbrowse&journal_id=1079991.

Aydinalp, M., Ugursal, V.I. & Fung, A.S., 2004. Modeling of the appliance, lighting, and space- cooling energy consumptions in the residential sector using neural networks. *Applied Energy*, 71, pp.87–110.

Azadeh, A. & Faiz, Z.S., 2011. A meta-heuristic framework for forecasting household electricity consumption. *Applied Soft Computing Journal*, 11(1), pp.614–620.

Beccali, M. et al., 2008. Short-term prediction of household electricity consumption: Assessing weather sensitivity in a Mediterranean area. *Renewable and Sustainable Energy Reviews*, 12(8), pp.2040–2065.

Bianco, V., Manca, O. & Nardini, S., 2009. Electricity consumption forecasting in Italy using linear regression models. *Energy*, 34(9), pp.1413–1421. Available at: http://dx.doi.org/10.1016/j.energy.2009.06.034.

Boser, B.E., Guyon, I.M. & Vapnik, V.N., 1992. A Training Algorithm for Optimal Margin Classifiers. *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory*, pp.144–152. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.21.3818.

Boulaire, F. et al., 2014. Statistical modelling of district-level residential electricity use in NSW, Australia. *Sustainability Science*, 9(1), pp.77–88.

ACCEPTED MANUSCRIPT

Box, G.E.P., Jenkins, G.M. & Reinsel, G.C., 2008. *Time series analysis: Forecasting and Control*, Available at: http://books.google.com/books?id=aY0QAQAAIAAJ&q=inauthor:Wei+Time+series+analysis+inpublisher:Addison-Wesley++Publishing&dq=inauthor:Wei+Time+series+analysis+inpublisher:Addison-Wesley++Publishing&ie=ISO-8859-1&cd=1&source=gbs_gdata LK - Preview%3Cspan sty.

Capasso, A. et al., 1994. A bottom-up approach to residential load modeling. , 9(2), pp.957–964.

Cheng, V. & Steemers, K., 2011. Modelling domestic energy consumption at district scale: A tool to support national and local energy policies. *Environmental Modelling and Software*, 26(10), pp.1186–1198. Available at: http://dx.doi.org/10.1016/j.envsoft.2011.04.005.

Cherkassky, V., 1997. The Nature Of Statistical Learning Theory. *IEEE Transactions on Neural Networks*, 8(6), pp.1564–1564. Available at: http://portal.acm.org/citation.cfm?id=211359.

Chujai, P., Kerdprasop, N. & Kerdprasop, K., 2013. Time Series Analysis of Household Electric Consumption with ARIMA and ARMA Models. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*.

Ciabattoni, L. et al., 2013. A Fuzzy Logic tool for household electrical consumption modeling. In *IECON Proceedings (Industrial Electronics Conference)*. pp. 8022–8027.

Cortes, C. & Vapnik, V., 1995. Support-Vector Networks. *Machine Learning*, 20(3), pp.273–297.

ACCEPTED MANUSCRIPT

Crone, S.F., Guajardo, J. & Weber, R., 2006. The impact of preprocessing on support vector regression and neural networks in time series prediction. In *Proceedings of the 2006 International Conference on Data Mining, DMIN*. pp. 37–44.

Dilaver, Z. & Hunt, L.C., 2011. Modelling and forecasting Turkish residential electricity demand. *Energy Policy*, 39(6), pp.3117–3127. Available at: http://dx.doi.org/10.1016/j.enpol.2011.02.059.

Eppi, 2007. *EPPI-Centre Methods form Conducting systematic review*,

Fan, H., MacGill, I.F. & Sproul, A.B., 2015. Statistical analysis of driving factors of residential energy demand in the greater Sydney region, Australia. *Energy and Buildings*, 105, pp.9–25. Available at: http://www.sciencedirect.com/science/article/pii/S0378778815301419.

Fan, J.Y., Mcdonald, J.D. & Member, S., 1994. A real-time implementation of short-term load forecasting for distribution power systems. *IEEE Transactions on Power Systems*, 9(2), pp.988–994.

Farzana, S. et al., 2014. Multi-model prediction and simulation of residential building energy in urban areas of Chongqing, South West China. *Energy and Buildings*, 81, pp.161–169. Available at: http://dx.doi.org/10.1016/j.enbuild.2014.06.007.

Filik, Ü.B., Gerek, Ö.N. & Kurban, M., 2011. A novel modeling approach for hourly forecasting of long-term electric energy demand. *Energy Conversion and Management*, 52(1), pp.199–211.

Fischer, D., Hartl, A. & Wille-Haussmann, B., 2015. Model for electric load profiles with high time resolution for German households. *Energy and Buildings*, 92, pp.170–179. Available at: http://dx.doi.org/10.1016/j.enbuild.2015.01.058.

ACCEPTED MANUSCRIPT

Garulli, A., Paoletti, S. & Vicino, A., 2015. Models and Techniques for Electric Load Forecasting in the Presence of Demand Response. *IEEE Transactions on Control Systems Technology*, 23(3), pp.1087–1097.

Generous, N. et al., 2014. Global Disease Monitoring and Forecasting with Wikipedia. *PLoS Computational Biology*, 10(11).

Ghods, L. & Kalantar, M., 2011. Different methods of long-term electric load demand forecasting; a comprehensive review. *Iranian Journal of Electrical and Electronic Engineering*, 7(4), pp.249–259.

Gonzales Chavez, S., Xiberta Bernat, J. & Llaneza Coalla, H., 1999. Forecasting of energy production and consumption in Asturias (northern Spain). *Energy*, 24(3), pp.183–198.

Gul, M., Qazi, S.A. & Qureshi, W.A., 2011. Incorporating economic and demographic variablesfor forecasting electricity consumption in Pakistan. In *2011 2nd International Conference on Electric Power and Energy Conversion Systems (EPECS)*. IEEE, pp. 1–5. Available at: http://ieeexplore.ieee.org/document/6126852/.

Hernandez, L. et al., 2014. Artificial neural networks for short-term load forecasting in microgrids environment. , 75, pp.252–264.

Hernandez, L. et al., 2013. Improved Short-Term Load Forecasting Based on Two-Stage Predictions with Artificial Neural Networks in a Microgrid Environment. , pp.4489–4507.

Hernández, L. et al., 2014. Artificial Neural Network for Short-Term Load Forecasting in Distribution Systems. *Energies*, 7(3), pp.1576–1598. Available at: http://www.mdpi.com/1996-1073/7/3/1576/.

Hernández, L. et al., 2012. Classification and Clustering of Electricity Demand Patterns in

ACCEPTED MANUSCRIPT

Industrial Parks. , pp.5215–5228.

Higgins JPT, Green S, editors. ed., 2006. *Cochrane Handbook for Systematic Reviews of Interventions 4.2.6* The Cochra., Available at: http://handbook.cochrane.org.

Hippert, H.S., Pedreira, C.E. & Souza, R.C., 2001. Neural networks for short-term load forecasting: a review and evaluation. *IEEE Transactions on Power Systems*, 16(1), pp.44–55.

Ho, S.., Xie, M. & Goh, T.., 2002. A comparative study of neural network and Box-Jenkins ARIMA modeling in time series prediction. *Computers & Industrial Engineering*, 42(2–4), pp.371–375. Available at: http://www.sciencedirect.com/science/article/pii/S0360835202000360.

Hoffman, A.J., 1998. Peak demand control in commercial buildings with target peak adjustment based on load forecasting. In *Proceedings of the 1998 IEEE International Conference on Control Applications (Cat. No.98CH36104)*. IEEE, pp. 1292–1296. Available at: http://ieeexplore.ieee.org/document/721669/.

Hsiao, Y.-H., 2015. Household Electricity Demand Forecast Based on Context Information and User Daily Schedule Analysis From Meter Data. *IEEE Transactions on Industrial Informatics*, 11(1), pp.33–43. Available at: http://ieeexplore.ieee.org/document/6926785/.

Huang, J., Li, Y.-F. & Xie, M., 2015. An empirical analysis of data preprocessing for machine learning-based software cost estimation. *Information and Software Technology*, 67, pp.108–127. Available at: http://www.sciencedirect.com/science/article/pii/S0950584915001275.

Hyndman, R.J. & Athanasopoulos, G., 2014. Summary for Policymakers. In

ACCEPTED MANUSCRIPT

Intergovernmental Panel on Climate Change, ed. *Climate Change 2013 - The Physical Science Basis*. Cambridge: Cambridge University Press, pp. 1–30. Available at: http://ebooks.cambridge.org/ref/id/CBO9781107415324A009.

Inglesi, R., 2010. Aggregate electricity demand in South Africa: Conditional forecasts to 2030. *Applied Energy*, 87(1), pp.197–204. Available at: http://dx.doi.org/10.1016/j.apenergy.2009.08.017.

Jurado, S. et al., 2015. Hybrid methodologies for electricity load forecasting: Entropy-based feature selection with machine learning and soft computing techniques. *Energy*, 86, pp.276–291.

Koprinska, I., Rana, M. & Agelidis, V.G., 2011. Yearly and seasonal models for electricity load forecasting. In *The 2011 International Joint Conference on Neural Networks*. IEEE, pp. 1474–1481. Available at: http://ieeexplore.ieee.org/document/6033398/.

Landau, L.J. & Taylor, J.G., 1998. *Concepts for Neural Networks* L. J. Landau & J. G. Taylor, eds., London: Springer London. Available at: http://link.springer.com/10.1007/978-1-4471-3427-5.

Makridakis, S. & Hibon, M., 2000. The M3-Competition: results, conclusions and implications. *International Journal of Forecasting*, 16(4), pp.451–476.

Marvuglia, A. & Messineo, A., 2012. Using recurrent artificial neural networks to forecast household electricity consumption. *Energy Procedia*, 14, pp.45–55. Available at: http://dx.doi.org/10.1016/j.egypro.2011.12.895.

Massana, J. et al., 2015. Short-term load forecasting in a non-residential building contrasting models and attributes. *Energy and Buildings*, 92, pp.322–330.

Mathieu, J.L. et al., 2011. Quantifying Changes in Building Electricity Use, With Application

ACCEPTED MANUSCRIPT

to Demand Response. *IEEE Transactions on Smart Grid*, 2(3), pp.507–518. Available at: http://ieeexplore.ieee.org/document/5772947/.

McLoughlin, F., Duffy, A. & Conlon, M., 2012. Characterising domestic electricity consumption patterns by dwelling and occupant socio-economic variables: An Irish case study. *Energy and Buildings*, 48(July 2009), pp.240–248. Available at: http://dx.doi.org/10.1016/j.enbuild.2012.01.037.

McLoughlin, F., Duffy, A. & Conlon, M., 2013. Evaluation of time series techniques to characterise domestic electricity demand. *Energy*, 50(1), pp.120–130. Available at: http://dx.doi.org/10.1016/j.energy.2012.11.048.

Mena, R. et al., 2014. A prediction model based on neural networks for the energy consumption of a bioclimatic building. *Energy and Buildings*, 82, pp.142–155. Available at: http://dx.doi.org/10.1016/j.enbuild.2014.06.052.

Mohandes, M., 2002. Support vector machines for short-term electrical load forecasting. *International Journal of Energy Research*, 26(4), pp.335–345. Available at: http://doi.wiley.com/10.1002/er.787.

Newsham, G.R. & Birt, B.J., 2010. Building-level occupancy data to improve ARIMA-based electricity use forecasts. In *Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building - BuildSys '10*. New York, New York, USA: ACM Press, p. 13. Available at: http://portal.acm.org/citation.cfm?doid=1878431.1878435.

Pantic, M., 2000. Introduction to Machine Learning & Case-Based Reasoning. In *Machine Learning*.

Parlos, A.G. et al., 1996. Development of an intelligent long-term electric load forecasting

ACCEPTED MANUSCRIPT

system. In *Proceedings of International Conference on Intelligent System Application to Power Systems*. IEEE, pp. 288–292. Available at: http://ieeexplore.ieee.org/document/501084/.

Platon, R., Dehkordi, V.R. & Martel, J., 2015. Hourly prediction of a building’s electricity consumption using case-based reasoning, artificial neural networks and principal component analysis. *Energy and Buildings*, 92, pp.10–18. Available at: http://dx.doi.org/10.1016/j.enbuild.2015.01.047.

Richardson, I. et al., 2010. Domestic electricity use: A high-resolution energy demand model. *Energy and Buildings*, 42(10), pp.1878–1887. Available at: http://dx.doi.org/10.1016/j.enbuild.2010.05.023.

Richardson, I. et al., 2009. Domestic lighting: A high-resolution energy demand model. *Energy and Buildings*, 41(7), pp.781–789.

Righi, A.W., Saurin, T.A. & Wachs, P., 2015. A systematic literature review of resilience engineering: Research areas and a research agenda proposal. *Reliability Engineering & System Safety*, 141, pp.142–152. Available at: http://www.sciencedirect.com/science/article/pii/S0951832015000654.

Song, Y. & Aggarwal, R., 1998a. Artificial neural networks in power systems. Part 2: Types of artificial neural networks. *Power Engineering Journal*, 12(1), pp.41–47. Available at: http://digital-library.theiet.org/content/journals/10.1049/pe_19980110.

Song, Y. & Aggarwal, R., 1998b. Artificial neural networks in power systems. Part 3: Examples of applications in power systems. *Power Engineering Journal*, 12(6), pp.279–287. Available at: http://digital-library.theiet.org/content/journals/10.1049/pe_19980609.

Suhartono & Subanar, 2006. The Effect of Decomposition Method As Data Preprocessing on

ACCEPTED MANUSCRIPT

Neural Networks Model for forecasting trend and seasonal time series. *JURNAL TEKNIK INDUSTRI*, 8(2), pp.156–164.

Swan, L.G., Ugursal, V.I. & Beausoleil-Morrison, I., 2011. Occupant related household energy consumption in Canada: Estimation using a bottom-up neural-network technique. *Energy and Buildings*, 43(2–3), pp.326–337.

Tranfield, D., Denyer, D. & Smart, P., 2003. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management*, 14, pp.207–222.

Twanabasu, S.R. & Bremdal, B.A., 2013. Load forecasting in a smart grid oriented building. In *22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013)*. Institution of Engineering and Technology, pp. 0907–0907. Available at: http://digital-library.theiet.org/content/conferences/10.1049/cp.2013.0997.

Wang, J.C., 2012. A study on the energy performance of hotel buildings in Taiwan. *Energy and Buildings*, 49, pp.268–275.

Wehenkel, L., 1997. Machine learning approaches to power-system security assessment. *IEEE Expert*, 12(5), pp.60–72. Available at: http://ieeexplore.ieee.org/ielx3/64/13513/00621229.pdf?tp=&arnumber=621229&isnumber=13513.

Widen, J. & Wackelgard, E., 2010. A high-resolution stochastic model of domestic activity patterns and electricity demand. *Applied Energy*, 87(6), pp.1880–1892. Available at: http://dx.doi.org/10.1016/j.apenergy.2009.11.006.

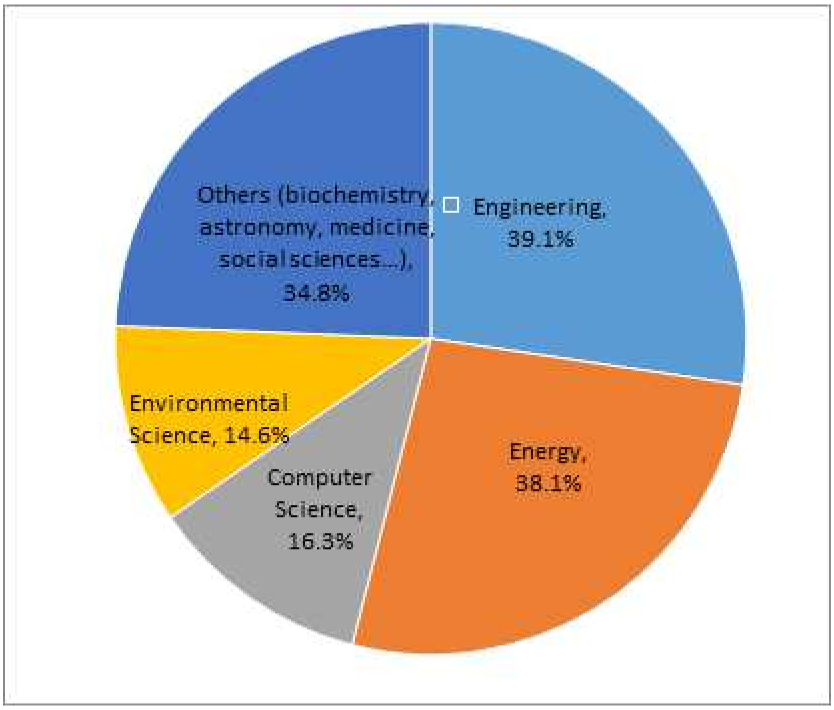
Yoo, J. & Hur, K., 2013. Load Forecast Model Switching Scheme for Improved Robustnessto Changes in Building Energy Consumption Patterns. *Energies*, 6(3), pp.1329–1343.

ACCEPTED MANUSCRIPT

Available at: http://www.mdpi.com/1996-1073/6/3/1329/.

Zahedi, G. et al., 2013. Electricity demand estimation using an adaptive neuro-fuzzy network: A case study from the Ontario province - Canada. *Energy*, 49(1), pp.323–328. Available at: http://dx.doi.org/10.1016/j.energy.2012.10.019.

Zhai, Y., 2005. *Time series forecasting competition among three sophisticated paradigms*. University of North Carolina at Wilmington.



ACCEPTED MANUSCRIPT

Figure 1 Fields' distribution through the papers

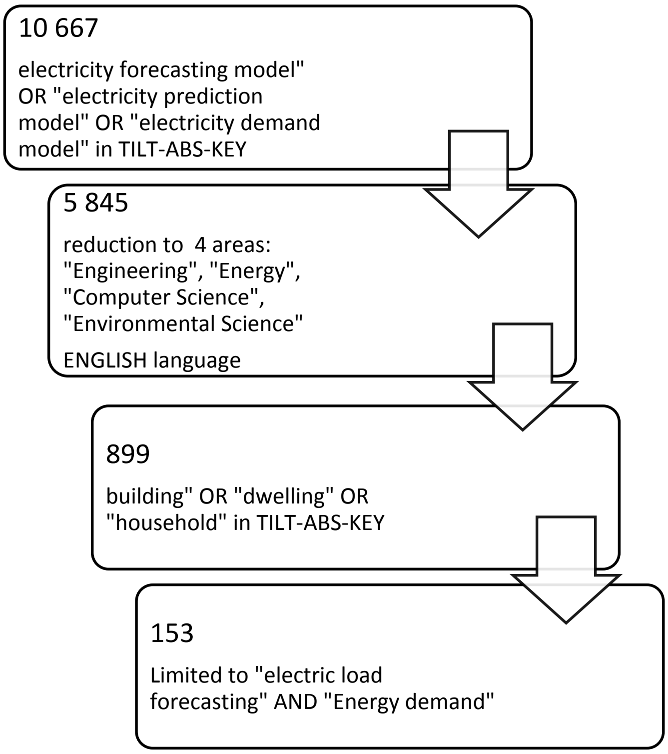
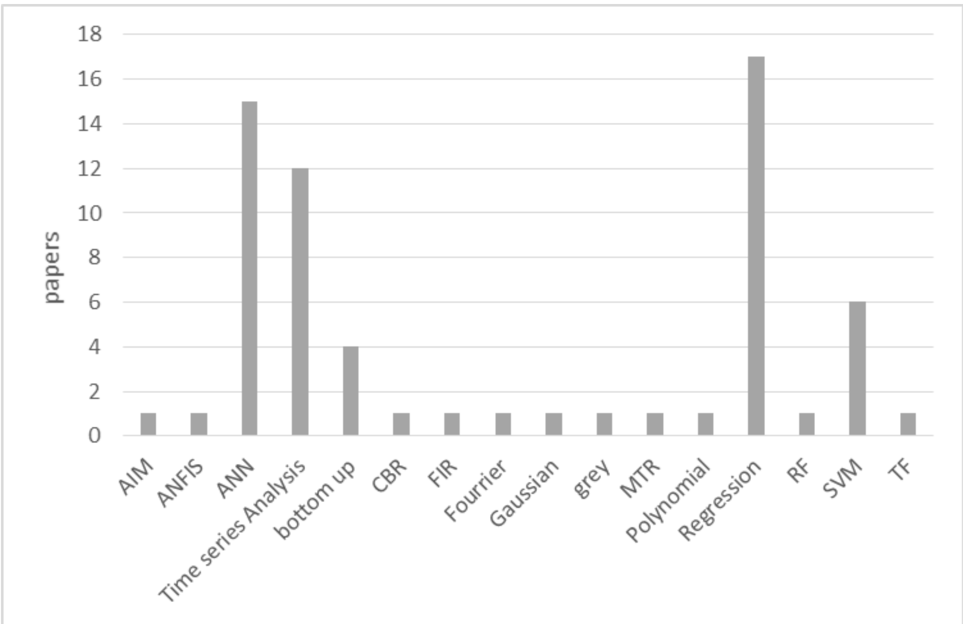


Figure 2 Selection procedure



ACCEPTED MANUSCRIPT

Figure 3 Classified forecasting models distribution

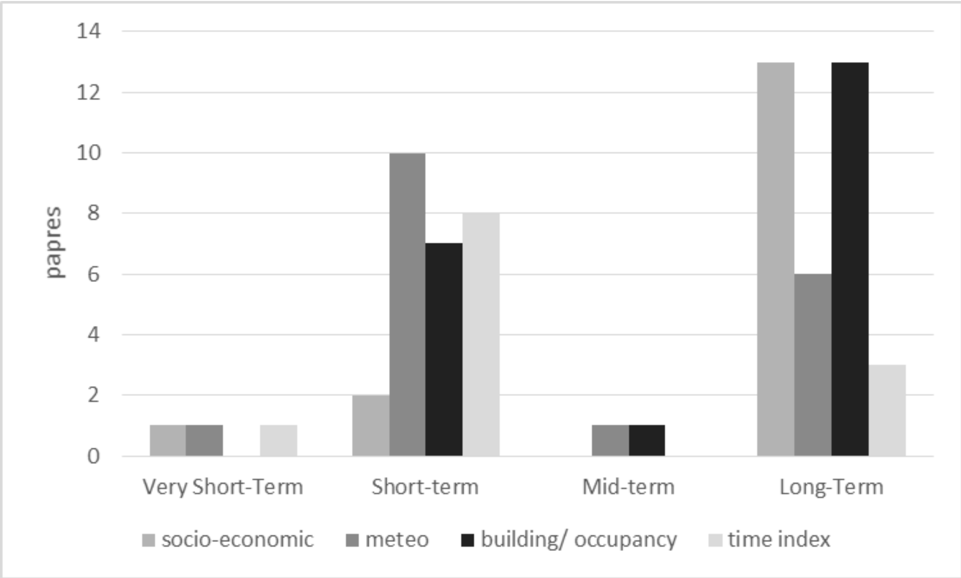


Figure 4 Input distribution depending of the time horizon

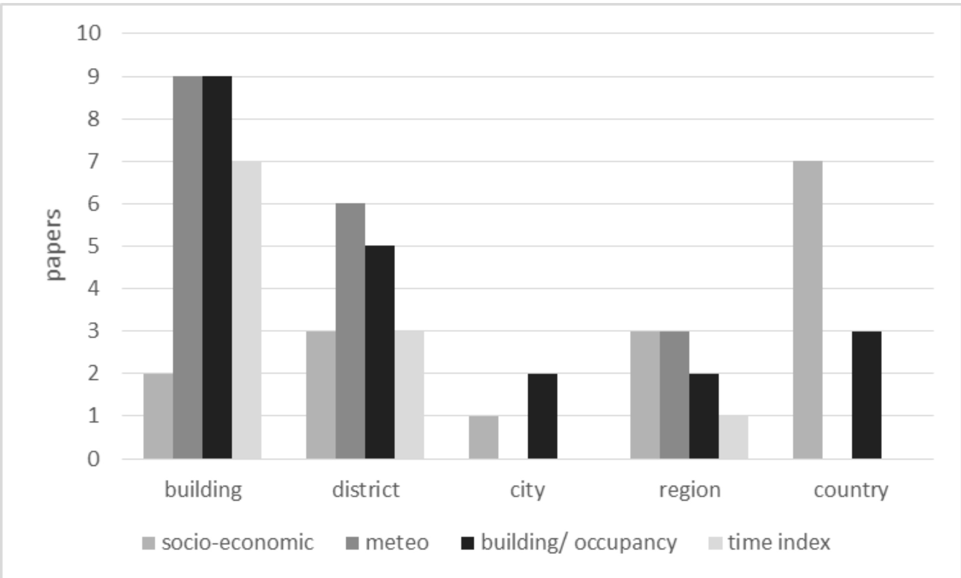


Figure 5 Input distribution depending of the scale

ACCEPTED MANUSCRIPT

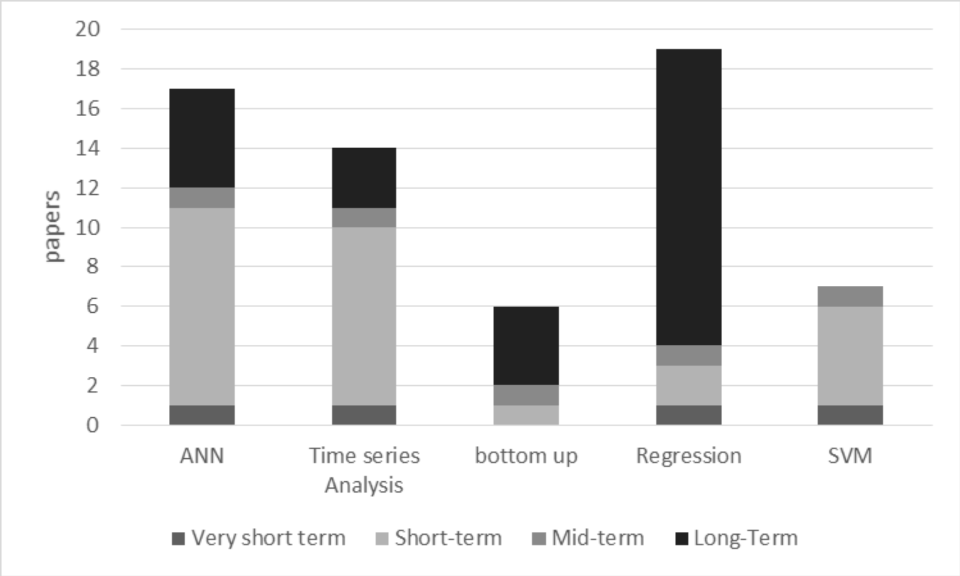


Figure 6 Models vs time horizon distribution

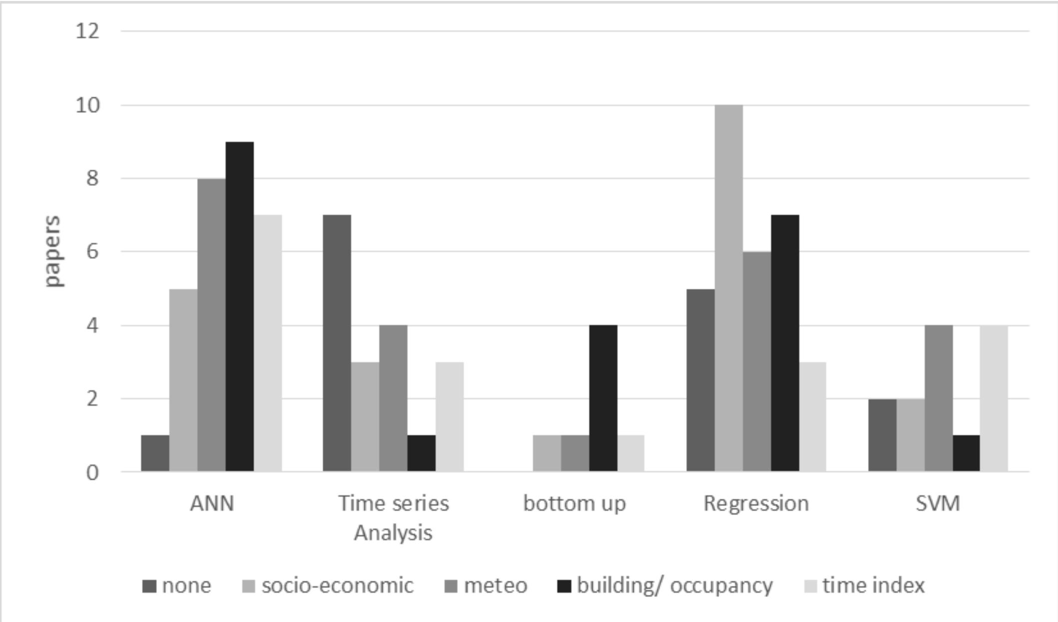


Figure 7 Models vs inputs distribution

ACCEPTED MANUSCRIPT

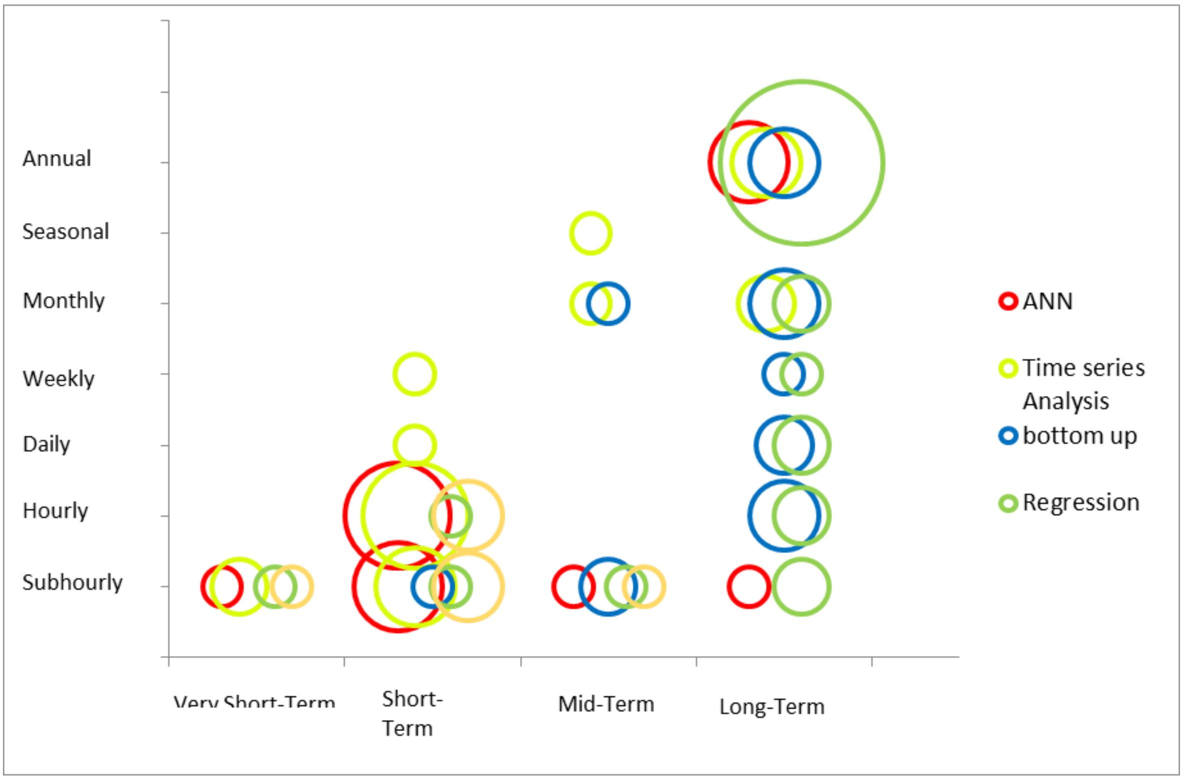
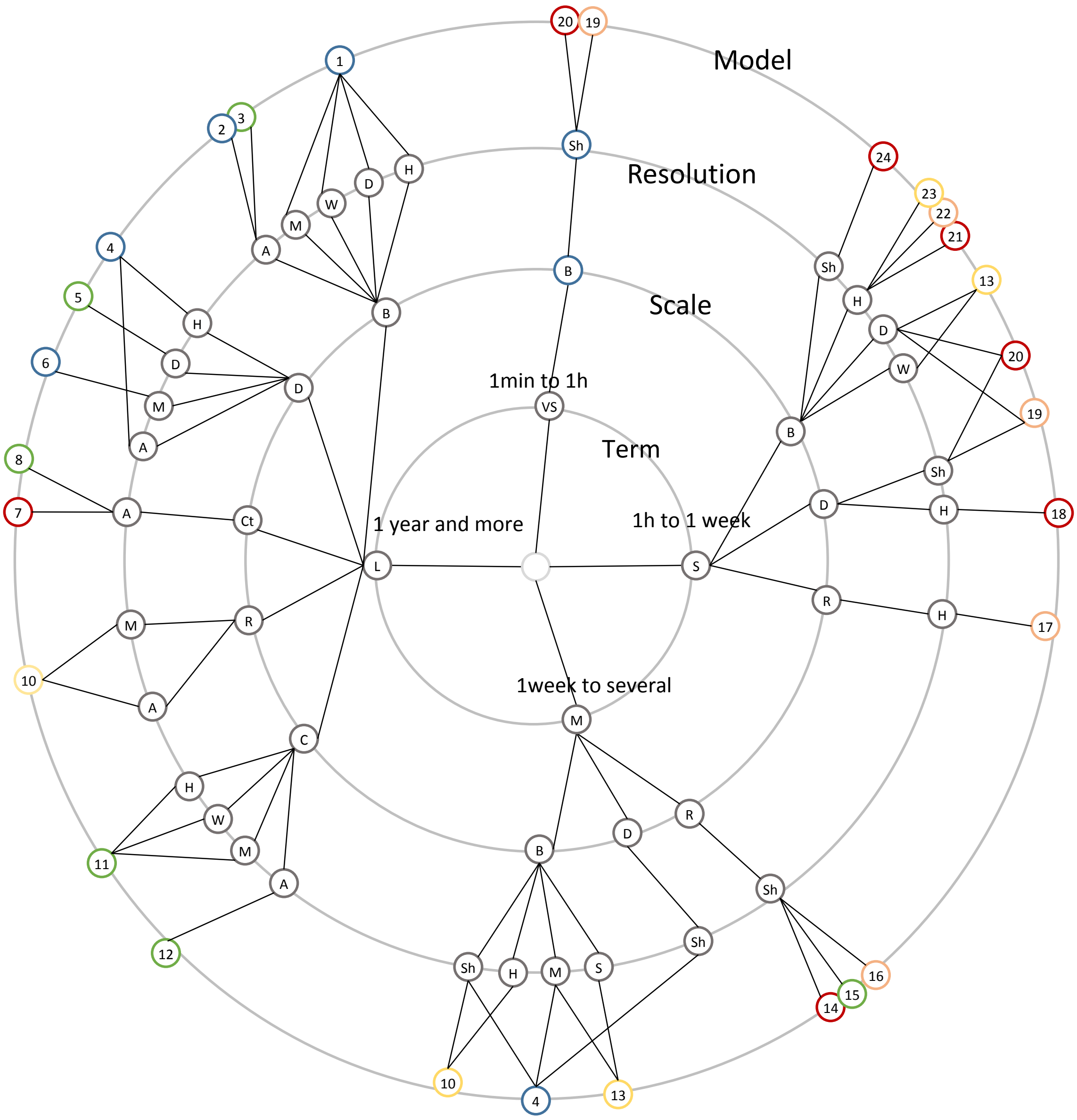


Figure 8 Models distribution by time horizon and resolution

ACCEPTED MANUSCRIPT



Term: Very-Short = VS; Short = S; Mid = M; Long = L
Scale: Building = B; District = D; City = Ct; Region = R; Country = C
Resolution: Subhourly = Sh; Hourly = H; Daily = D; Weekly = W; Monthly = M; Seasonal = S; Annual = A

Bottom up   Regression   ANN   Time Series   SVM

Figure 9 Study taxonomy

ACCEPTED MANUSCRIPT

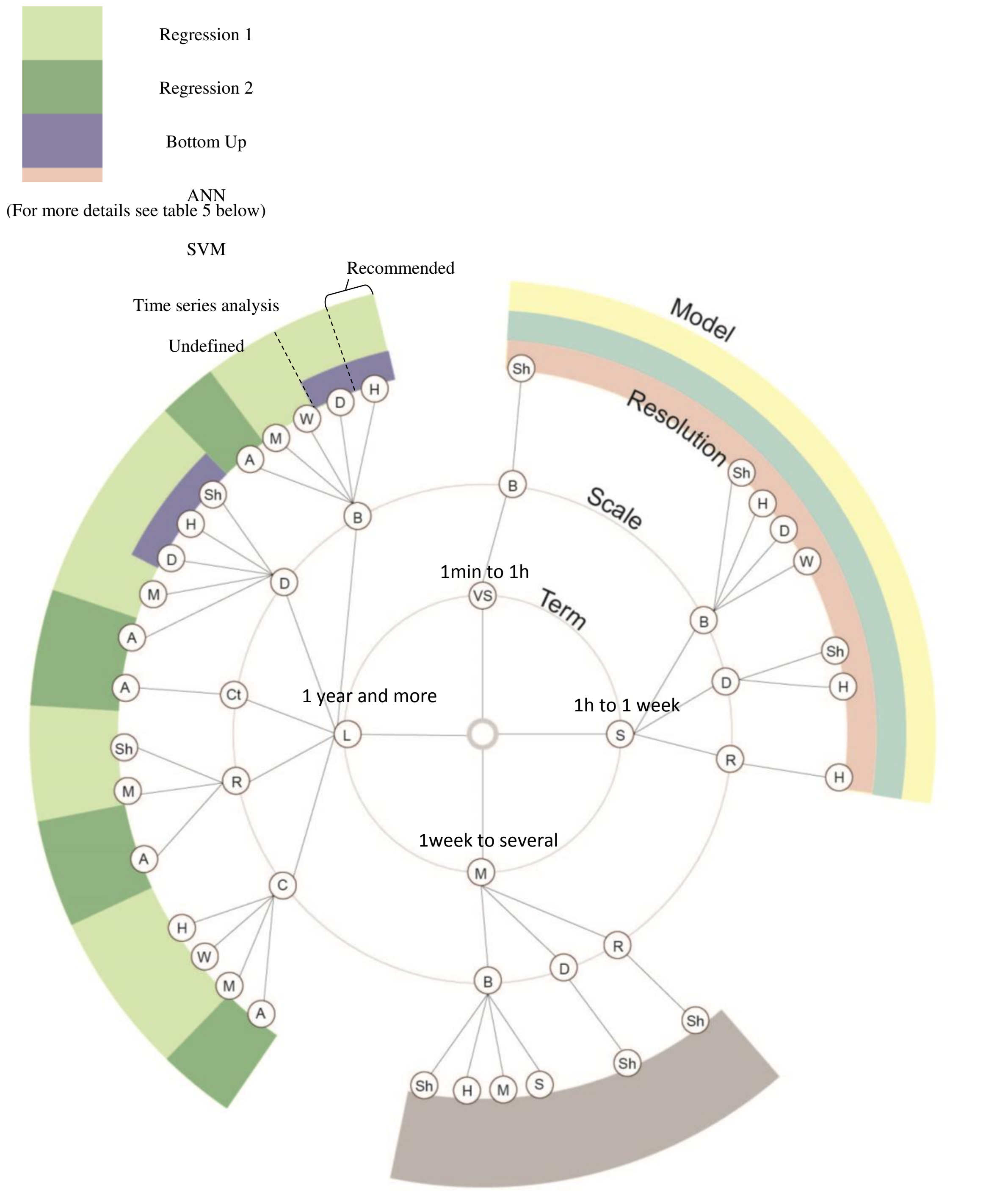


Figure 10 General taxonomy

ACCEPTED MANUSCRIPT

Table 1 Cases characteristics

Frame	Describe the context of the study and give an overview of the study purposes.
Location	Country of the case.
Scale	Scale of the study, from a single building to an entire country. Size of the sample are taken into account.
Term	From very short (1 min ahead) to very long term (several years ahead), gives the timeframe of the case.
Time resolution	Gives the time step considerate in the forecast: every minute, hours, day, years…
Inputs	Inputs implemented in the forecasting model.
Historical Data	Gives the length of the data sample used for the prediction as well as their origin (meters, statistical…).
Pre-processing	Indicates if the data have been pre-processed before being introduced into the forecasting model and which type of pre-process have been done.
Forecasting model	Gives the forecasting model employed
Error	Gives the type of error measured (CV(RMSE), MAPE, RMSE…).

ACCEPTED MANUSCRIPT

Table 2 Term distribution through the reviewed papers and cases

Term	nb studies	nb papers	Distribution percentage
Very Short-Term	5	1	2.6%
Short-term	40	17	43.6%
Mid-term	9	3	7.7%
Long-Term	58	24	61.5%

Table 3 Inputs distribution through the papers and cases

Exogenous Input	nb studies	nb paper	Distribution percentage
None	36	12	30.8 %
Socio-economic	42	15	38.5%
Environmental	41	16	41.0%
Building	44	19	48.7%
Time index	33	11	28.2%

ACCEPTED MANUSCRIPT

Table 4 Taxonomy references

#	Model	Input resolution	Exogenous Inputs data	Pre-process recommended	Reference
1	Bottom up	Subhourly	Building & time index	Smoothing	(Ciabattoni et al. 2013; Fischer et al. 2015)
2	Bottom up	Subhourly	Building & socio-economic	None	(Fischer et al. 2015)
3	Regression	Monthly	Building	Dependency & significance	(Wang 2012)
4	Bottom up	Subhourly	Building & environmental	Smoothing	(Widen & Wackelgard 2010; Richardson et al. 2009)
5	Regression	Daily	Socio-economic, environmental, building	Dependency & significance	(Fan et al. 2015)
6	Bottom up	Subhourly	Building	None	(Richardson et al. 2009)
7	ANN	Annual	Socio-economic & building	None	(Aydinalp et al. 2004; Farzana et al. 2014)
8	Regression	Annual	Socio-economic & building	None	(Farzana et al. 2014)
9	ANN	Subhourly	None	Clustering	(Koprinska et al. 2011)
10	TSA	Monthly	None	None	(Abdel-aal & Al-Garni 1997; Gonzales Chavez et al. 1999)
11	Regression	Hourly	None	Smoothing, Clustering	(Filik et al. 2011; Al-Hamadi & Soliman 2005)
12	Regression	Annual	Socio-economic	Dependency & Stationarity eventually	(Filik et al. 2011; Azadeh & Faiz 2011; Bianco et al. 2009; Gul et al. 2011; Dilaver & Hunt 2011; Al-Ghandoor et al. 2009)
13	TSA	Subhourly	None	Smoothing, Clustering	(Chujai et al. 2013)
14	ANN	Subhourly	None	Clustering	(Koprinska et al. 2011)
15	Regression	Subhourly	None	Clustering	(Koprinska et al. 2011)
16	SVM	Subhourly	None	Clustering	(Koprinska et al. 2011)
17	SVM	Hourly	None	Smoothing, Clustering	(Mohandes 2002)
18	ANN	Hourly	Environmental, building & time index	Clustering	(Hernández et al. 2014; Marvuglia & Messineo 2012)
19	SVM	Subhourly	Socio-economic, environmental, time index	Smoothing, Clustering	(Garulli et al. 2015; Hsiao 2015)

ACCEPTED MANUSCRIPT

20	ANN	Subhourly	Socio-economic, environmental, time index	Smoothing, Clustering	(Garulli et al. 2015; Hsiao 2015)
21	ANN	Hourly	Environmental, building & time index	Smoothing, Dependency & significance	(Platon et al. 2015; Jurado et al. 2015; Twanabasu & Bremdal 2013; Massana et al. 2015; Beccali et al. 2008)
22	SVM	Hourly	Environmental, building & time index	Smoothing eventually	(Twanabasu & Bremdal 2013; Massana et al. 2015)
23	TSA	Hourly	Environmental	Smoothing	(Newsham & Birt 2010; Jurado et al. 2015; Yoo & Hur 2013; Twanabasu & Bremdal 2013)
24	ANN	Subhourly	Environmental, building & time index	Dependency & significance	(Mena et al. 2014)

Table 5 General taxonomy references

Colour	Model	Inputs resolution	Inputs	Pre-process
	Regression	Subhourly to Hourly	Building data can eventually be introduced for better performance.	Smoothing high-resolution dataset is recommended. Clustering dataset in seasonal pattern can eventually be done to improve performance.
	Regression	Hourly to Annual	Socio-economic data are often introduced	Dependency & significance in order to lower the amount of input data
	Bottom Up	Subhourly	Building data are always introduced due to the nature of the model. Environmental data and time index can eventually be introduced for better performance.	Smoothing high resolution data is recommended
	ANN	Subhourly to Hourly	Can be set with a large variety of data. Mainly environmental, building and time index. Time index often improve performance.	Dependency & significance in order to lower the amount of input data. Smoothing high-resolution data is recommended.
	SVM	Subhourly to Hourly	Can be set with a large variety of data. Mainly environmental, building and time	Smoothing high-resolution data is recommended.

ACCEPTED MANUSCRIPT

			index. Time index often improve performance.	Clustering dataset in seasonal pattern can eventually be done to improve performance.
	Time series analysis	Subhourly to Hourly	Environmental and time index data can eventually be introduced for better performance.	Smoothing high-resolution data is recommended. Clustering dataset in seasonal pattern can eventually be done to improve performance.
	The few amount of cases on mid-term forecast does not allow any generalisation on the model to employ. A large variety of unique models have been employed in the literature.			